

(18) ARO 11077.6-M

(19)

13

ARO Report 75-2

2 right for 75.

R - 2

(6)

PROCEEDINGS OF THE TWENTIETH CONFERENCE
ON THE DESIGN OF EXPERIMENTS IN ARMY
RESEARCH DEVELOPMENT AND TESTING.

PART II

Ranking and Selection Procedures.

(12) 23p.

(10) Robert E. / Bechhofer

(11) 1975

ADA019772



(15)

DAHC04-73-C-0008,
NO0014-67-A-0077-002

Approved for public release; distribution unlimited.
The findings in this report are not to be construed
as an official Department of the Army position, un-
less so designated by other authorized documents.

Sponsored by
The Army Mathematics Steering Committee
on Behalf of

DDC
RECEIVED
DEC 16 1975
D

THE CHIEF OF RESEARCH, DEVELOPMENT AND ACQUISITION

403133

ACQUISITION OF

NTIS

Whole Section



DDI

Self Section



UNANNOUNCED



JUSTIFICATION

BY

DISTRIBUTION/AVAILABILITY CODES

Disl.

AVAIL. No 4/ or SPECIAL

A		
---	--	--

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER ARO 11977.6-M	2. GOVT ACCESSION NO.	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (or Subtitle) RANKING AND SELECTION PROCEDURES		5. TYPE OF REPORT & PERIOD COVERED
		6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR Robert E. Bechhofer		8. CONTRACT OR GRANT NUMBER DAHCO4-73-C-0008, N00014-67-A-0077-0020
9. PERFORMING ORGANIZATION NAME AND ADDRESS Cornell University Ithaca, New York 14853		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS
11. CONTROLLING OFFICE NAME AND ADDRESS U. S. Army Research Office Post Office Box 12211 Research Triangle Park, NC 27709		12. REPORT DATE 1975
14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office)		13. NUMBER OF PAGES 21
		15. SECURITY CLASS. (of this report) Unclassified
		16. DECLASSIFICATION/DOWNGRADING SCHEDULE
16. DISTRIBUTION STATEMENT (of this Report) Approved for public release; distribution unlimited.		
17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report) D D C RECEIVED DEC 16 1975 D		
18. SUPPLEMENTARY NOTES The findings in this report are not to be construed as an official Department of the Army position, unless so designated by other authorized documents.		
19. KEY WORDS (Continue on reverse side if necessary and identify by block number) Statistics Experimental design Ranking Decision theory Selection		
20. ABSTRACT (Continue on reverse side if necessary and identify by block number) The purpose of this paper is to introduce the reader to ranking and selection procedures, to describe some of them and the philosophy underlying their use, and to discuss their properties.		

ADA019777

RANKING AND SELECTION PROCEDURES*

Robert E. Bechhofer
Department of Operations Research
Cornell University
Ithaca, New York 14853

1. Introduction

✓ In recent years statisticians have become increasingly concerned with the meaningful formulation and solution of certain multiple-decision problems which arise in experimentation. Thus, for example, when an experimenter conducts tests to compare the performances of several competing categories of items, his ultimate objective often is to select the category (or categories) which is (are) best, goodness being measured in terms of a particular parameter (e.g., the population mean or the population variance) associated with the random variable being observed. To accomplish this the experimenter requires a statistical decision procedure which will tell him how many observations to take, how to take these observations, and based on these observations which population(s) to choose; the decision procedure should have the property that the probability of an incorrect selection (or more generally, the risk or expected loss) is controlled at some specified level.

In response to the need for such decision procedures, research statisticians have been studying various possible appropriate formulations of these problems, and have developed a body of statistical methodology to cope with them. The procedures have come to be referred to as ranking and selection procedures. The purpose of this paper is to introduce the reader to these procedures, to describe some of them and the philosophy underlying their use, and to discuss their properties.

In Section 2 we will pose the normal means problem, and use it as a vehicle for motivating some of the basic ideas. The two most commonly adopted formulations of ranking and selection problems, namely the so-called indifference-zone approach and the subset approach, will be described. The attributes of single-stage, two-stage, and sequential procedures devised for the normal means problem, under different assumptions concerning the population variances, will be assessed. In Section 3 we sketch some analogous results for the normal variances problem, and in Section 4 we mention results for parameters of other distributions.

*This research was supported in part by the U.S. Army Research Office-Durham under Contract DANC04-73-C-0008 and by the Office of Naval Research under Contract N00014-67-A-0077-0020.

The number of research papers written on subjects in this field is now vast; it is hoped that this brief introduction will stimulate the reader to explore the literature, and to apply the procedures where appropriate.

2. The normal means problem

A very important problem which arises frequently in applications is that of selecting the normal population which has the largest population mean. Thus, for example, the ordnance engineer might be conducting firing programs to compare the ballistic performance of different types of projectiles (in which case his objective might be to select that type which, on the average, travels the greatest distance), or the medical research worker might be studying the response of patients to different kinds of analgesic drugs (in which case his interest might lie in selecting that drug which produces, on the average, the longest period of time without pain), or the agronomist might be conducting field trials with different varieties of grain (in which case his purpose might be to select that variety which produces, on the average, the largest yield per acre). In all of these cases large values of the means are deemed to be desirable; however, in other cases small values of the means might be considered desirable. The procedures that we will describe can, with minor modifications, handle these latter cases as well.

In Sections 2.1 and 2.3.1 we shall state the statistical assumptions which underlie the procedures that have been developed. Then we shall describe several approaches to the selection problem.

2.1 Statistical assumptions

We shall assume that we have k sources Π_i ($1 \leq i \leq k$) of normally distributed data, the i th source having population mean μ_i and population variance σ_i^2 ; population Π_i ($1 \leq i \leq k$) should be thought of as being associated with the i th category. The μ_i are assumed to be unknown. Let $\mu_{[1]} \leq \mu_{[2]} \leq \dots \leq \mu_{[k]}$ denote the ranked values of the μ_i ; it is assumed that the pairing of the Π_i with the $\mu_{[j]}$ ($1 \leq i, j \leq k$) is completely unknown. Possible assumptions concerning the σ_i^2 ($1 \leq i \leq k$) will be discussed in Section 2.3.1. Throughout this paper x_{ij} ($1 \leq i \leq k, j=1, 2, \dots$) will denote the j th observation from Π_i , all observations being assumed independent.

2.2 Some formulations

The two most commonly used formulations of the selection problem are due to Bechhofer [1954] and Gupta [1956], [1965]; these are referred to as the indifference-zone approach and the subset approach, respectively. The approaches are described below.

2.2.1 The indifference-zone approach

The goal and probability requirement associated with the indifference-zone approach are:

Goal: "To select the population associated with $\mu_{[k]}$." (2.1)

It is assumed that prior to the start of experimentation the experimenter can specify two constants $\{\delta^*, P^*\}$ ($0 < \delta^* < \infty, 1/k < P^* < 1$) which are then incorporated into the following probability requirement:

Probability requirement:

$$\begin{aligned} \text{Prob}\{\text{Selecting the population associated with } \mu_{[k]}\} &\geq P^* \\ \text{whenever } \mu_{[k]} - \mu_{[k-1]} &\geq \delta^*. \end{aligned} \quad (2.2)$$

The experimenter then restricts consideration to procedures which guarantee (2.2). (In (2.2) the specified quantity δ^* can be thought of as the smallest difference "worth detecting" between the population mean of the "best" and "second best" population; P^* is specified strictly greater than $1/k$ since a probability of $1/k$ can be achieved by choosing one of the k populations at random.)

2.2.2 The subset approach

The goal and probability requirement associated with the subset approach are:

Goal: "To select a (non-empty) subset of the populations which contains the population associated with $\mu_{[k]}$." (2.3)

It is assumed that prior to the start of experimentation the experimenter can specify a constant $\{P^*\}$ ($1/k < P^* < 1$) which is then incorporated into the following probability requirement:

Probability requirement:

Prob{Selected subset contains the population associated with $\mu_{[k]}$ } $\geq P^*$
regardless of the values of the μ_i ($1 \leq i \leq k$). (2.4)

The experimenter then restricts consideration to procedures which guarantee (2.4).

Remark 1: It is to be noted that the experimenter plans his experiment assuming that the population means are not all equal; this is a very reasonable assumption in almost all real-life situations. He is interested in identifying the "best" population -- in this case the population with the largest population mean. Goal (2.1) leads to a k-decision problem since the experimenter must choose one of the k populations based on the outcome of his experiment (i.e., his possible decisions are: Π_1 is best, or Π_2 is best, or ..., or Π_k is best). Similarly, goal (2.3) leads to a $(2^k - 1)$ -decision problem since the experimenter must choose one of the $2^k - 1$ non-empty subsets of the k populations based on the outcome of his experiment (e.g., for $k=3$ his possible decisions are: only Π_1 is in the subset, only Π_2 is in the subset, only Π_3 is in the subset, Π_1 and Π_2 are in the subset, Π_1 and Π_3 are in the subset, Π_2 and Π_3 are in the subset, Π_1 and Π_2 and Π_3 are in the subset). These multi-decision approaches are in marked contrast to the classical 2-decision test-of-homogeneity approach afforded by the Analysis of Variance; in that approach the experimenter tests the (usually completely unrealistic) hypothesis that the k population means are equal, and decides based on the outcome of the experiment either to accept the hypothesis or to reject the hypothesis.

Remark 2: As noted above, goal (2.1) leads to a k -decision problem. However, depending on the practical situation under consideration, the experimenter can, using the indifference-zone approach, pose more general goals. For example, he may wish to select the t ($1 \leq t \leq k-1$) best populations with regard to order, or he may wish to select the t ($1 \leq t \leq k-1$) best populations without regard to order, t being fixed before the start of experimentation. (Both goals reduce to (2.1) when $t=1$.) These more general goals lead to a $[k!/(k-t)!]$ -decision problem and a $[k!/t!(k-t)!]$ -decision problem, respectively. Such general goals and others are discussed in Bechhofer [1954] and Mahamunulu [1967].

Remark 3: For goal (2.1) and the indifference-zone approach, the experimenter always ends up by selecting a single population. For goal (2.3) and the subset approach, the experimenter ends up by selecting 1 or 2 or ... or k populations, depending on the outcome of the experiment; thus for this latter approach the number of populations in the selected subset is a random variable.

2.2.3 Other approaches

Santner [1975] has proposed a restricted subset approach in which the experimenter selects 1 or 2 or ... or c populations, depending on the outcome of the experiment, where c ($1 \leq c \leq k$) is decided on and fixed before the start of experimentation; his approach can be regarded as bridging the indifference-zone and subset approaches since if $c=1$ his approach reduces to the indifference-zone approach while if $c=k$ it reduces to the subset approach. Other approaches in which more general "loss functions" are used have been proposed by Somerville [1954] and Fairweather [1968]. An approach in which the μ_i are assumed to have prior distributions has been considered by Dunnett [1960] while a similar idea from a Bayesian point of view has been proposed by Raiffa and Schlaiffer [1961] and Deely and Gupta [1968]. However, for brevity we will not discuss these or other approaches.

2.3 Assumptions concerning the variances

2.3.1 Possible assumptions

In order to devise procedures which will guarantee (2.2) or (2.4) for the normal means problem, it is necessary to make an assumption concerning the values of the σ_i^2 ($1 \leq i \leq k$). Which assumption it is appropriate for the experimenter to make in any particular practical situation depends on the information available to him at the time that he plans his experiment. The four most common assumptions are that:

- a) The values of the σ_i^2 ($1 \leq i \leq k$) are known, and all are equal to σ^2 (say). (2.5a)
- b) The values of the σ_i^2 ($1 \leq i \leq k$) are known, but not all are equal. (2.5b)
- c) The values of the σ_i^2 ($1 \leq i \leq k$) are unknown, but it is known that they have a common value σ^2 (say). (2.5c)
- d) The values of the σ_i^2 ($1 \leq i \leq k$) are completely unknown. (2.5d)

2.3.2 The variance assumption and associated procedures

Once the experimenter has adopted one of these assumptions he then must choose a selection procedure which was derived under that particular assumption.

Thus, for example, if he wishes to guarantee (2.2) and adopts assumption (2.5a) or (2.5b), then he can use a single-stage procedure (Bechhofer [1954]), a two-stage procedure (Alam [1970] or (Tamhane [1975])), an open sequential procedure without elimination (Bechhofer, Kiefer, Sobel [1968]), or a closed sequential procedure with elimination (Paulson [1964]). If he wishes to guarantee (2.2) and adopts assumption (2.5c), then he cannot use a single-stage procedure (see Dudewicz [1971]) although he can use a two-stage procedure (Bechhofer, Dunnett, Sobel [1954]) or a sequential procedure (Paulson [1964]); similarly, if he wishes to guarantee (2.2) and adopts assumption (2.5d), then he cannot use a single-stage procedure although he can use a two-stage procedure (Dudewicz and Dalal [1971] or Rinott [1974]). Finally, if the experimenter wishes to guarantee (2.4), and he adopts assumption (2.5a) or (2.5c), then he can use a single-stage procedure (Gupta [1956], [1965]).

When the experimenter has adopted a particular assumption and as a consequence has the option of choosing among several competing procedures, each one of which will guarantee his probability requirement, he then chooses one of these procedures on the basis of various possible operational or cost criteria. An indication of such criteria will be given in our later discussion. In the next section we shall describe certain selection procedures. Our emphasis will be on procedures which can be used with the indifference-zone approach to guarantee (2.2).

2.4 Procedures for use with the indifference-zone approach under the assumption of common known variance

In this section we shall describe three procedures, each one of which will guarantee (2.2) when assumption (2.5a) is made; minor modifications of these procedures will guarantee (2.2) when assumption (2.5b) is made. The procedures will be introduced in the order of their historical development, each being designed to afford different options to the experimenter.

2.4.1 Single-stage procedure

The easiest type of procedure to implement is a single-stage one. The following single-stage procedure was proposed by Bechhofer [1954]; constants c_{k,p^*} (see a), below) necessary to implement this procedure are given in Table I.

"a) Take a common number N of observations from each of the k populations where N is the smallest integer greater than or equal to $(c_{k,p^*}\sigma/\delta^*)^2$.

b) Calculate $\bar{X}_i = \sum_{j=1}^N X_{ij}/N$ ($1 \leq i \leq k$), and let $\bar{X}_{[1]} < \bar{X}_{[2]} < \dots < \bar{X}_{[k]}$ denote the ranked values of the \bar{X}_i . (2.6)

c) Select the population which yielded $\bar{X}_{[k]}$ as the one associated with $\mu_{[k]}$."

Note: The constants c_{k,p^*} are computed under the assumption that the μ_i ($1 \leq i \leq k$) are in the so-called least-favorable (LF) - configuration, i.e., $\mu_{[1]} = \mu_{[k-1]} = \mu_{[k]} - \delta^*$.

Table I
Values of c_{k,p^*}

p^*	k					
	2	3	4	5	7	10
0.99	3.2900	3.6173	3.7970	3.9196	4.0861	4.2456
0.95	2.3262	2.7101	2.9162	3.0552	3.2417	3.4182
0.90	1.8124	2.2302	2.4516	2.5997	2.7972	2.9829
0.80	1.1902	1.6524	1.8932	2.0528	2.2639	2.4608
0.60	0.3583	0.8852	1.1532	1.3287	1.5583	1.7700

The values in this table are abstracted from Table I of Bechhofer [1954] where values for other k and p^* are also given. Additional values for $k = n+1 = 2(1)51$ and $p^* = 0.99, 0.975, 0.95, 0.90, 0.75$ are contained in Table I of Gupta [1963]; Gupta's values must be multiplied by $\sqrt{2}$ in order to obtain the c_{k,p^*} - values required in (2.6).

2.4.2 Open-ended sequential procedure without elimination

The single-stage procedure of Section 2.4.1 is conservative in the sense that the constants c_{k,p^*} necessary to implement it are computed under the assumption that the population means are in the LF-configuration; however, it has been shown (Hall [1959]) that the probability requirement (2.2) cannot be guaranteed with a smaller N if the experimenter restricts consideration to single-stage procedures. If this restriction is eliminated, and multistage procedures are permitted, then certain gains can be achieved. What is desired is a multi-stage procedure which not only will guarantee the probability requirement (2.2) when the population means are in the LF-configuration, but also will require a smaller number of observations per population, on the average, than the N of (2.6) when the population means are in very favorable configurations--in particular when $(\mu_{[k]} - \mu_{[k-1]})/\sigma$ is large. The following sequential procedure, which possesses these attributes, was proposed by Bechhofer, Kiefer, and Sobel [1968], pp. 258-9, 264-7.

- "a) Take one observation from each of the k populations at each stage of experimentation. Let $\sum_{j=1}^m X_{ij}$ denote the cumulative sum from π_i ($1 \leq i \leq k$) at the m th stage of experimentation, and let $\sum_{j=1}^m X_{[1]j} < \sum_{j=1}^m X_{[2]j} < \dots < \sum_{j=1}^m X_{[k]j}$ denote the ranked values of the $\sum_{j=1}^m X_{ij}$.
- b) At the m th stage of experimentation ($m=1,2,\dots$) compute

$$Z_m = \sum_{i=1}^{k-1} \exp \left\{ -\frac{\delta^*}{\sigma} \frac{\sum_{j=1}^m X_{[k]j} - \sum_{j=1}^m X_{[i]j}}{\sigma} \right\}. \quad (2.7)$$

Then proceed as follows:

- i) If $Z_m \leq (1-p^*)/p^*$, stop experimentation and select the population which yielded $\sum_{j=1}^m X_{[k]j}$ as the one associated with $\mu_{[k]}$.

ii) If $Z_m > (1-P^*)/P^*$, take another observation from each of the k populations and compute Z_{m+1} .

Continue in this manner until the rule calls for stopping."

Remark 4: For (2.7) the observations are taken in vectors, each vector constituting a stage, there being one observation from each population in every vector. The number of stages (i.e., number of observations per population) necessary to terminate experimentation is a random variable. The expected number of stages to terminate experimentation has been shown (B-K-S [1968], Tables 12.8.2 and 12.8.3) to be less than N for many configurations of the μ_i ($1 \leq i \leq k$); in particular, if $(\mu_{[k]} - \mu_{[k-1]})/\sigma$ is large, then with high probability experimentation will cease after only a small number of stages. Regardless of the configuration of the μ_i ($1 \leq i \leq k$) experimentation will cease with probability one after a finite number of stages.

2.4.3 Closed sequential procedure with elimination

The sequential procedure of Section 2.4.2 has two possible drawbacks:

i) It is opened, i.e., before the start of experimentation it is not possible to give a finite upper bound on the number of stages to terminate experimentation, and ii) It does not eliminate "non-contending" populations, i.e., it continues to sample from populations which, based on observations obtained in the early stages of experimentation, would appear to be out of contention for being selected as "best." The following sequential procedure, which overcomes these drawbacks of (2.7), was proposed by Paulson [1964]; like (2.7) it guarantees the probability requirement (2.2) when the population means are in the LF-configuration, and also tends to cease experimentation early when the population means are in very favorable configurations:

For fixed λ ($0 < \lambda \leq \delta^*/2$) let $a_\lambda = [\sigma^2/(\delta^* - \lambda)] \log[(k-1)/(1-P^*)]$, and let W_λ = the largest integer less than a_λ/λ . Paulson's procedure is actually a family of procedures which depend on the choice of λ ; in Remark 9, below, we shall make some comments on the role of λ .

"Take one observation from each of the k populations at the first stage of experimentation. Eliminate from further consideration any population Π_i for which $a_\lambda - \lambda < \max_{1 \leq s \leq k} X_{s1} - X_{i1}$. If

all but one population is eliminated after the first stage, stop experimentation and select the remaining population as the one associated with $\mu_{[k]}$. Otherwise, go on to the second stage and take one observation from each population not yet eliminated. At stage m ($2 \leq m \leq W_\lambda$) take one observation from each population not eliminated after the $(m-1)$ st stage, and then eliminate from further consideration any remaining population Π_i for which

(2.8)

$$a_\lambda - m\lambda < \max_s \left\{ \sum_{j=1}^m x_{sj} \right\} - \sum_{j=1}^m x_{ij}$$

where the sums are only for populations left after the $(m-1)$ st stage. If all but one population is eliminated after the m th stage, stop experimentation and select the remaining population as the one associated with $\mu_{[k]}$; otherwise go on to the $(m+1)$ st stage. If more than one population remains after stage W_λ , terminate experimentation at the $(W_\lambda+1)$ st stage by selecting the remaining population with the largest sum of the $(W_\lambda+1)$ observations as the one associated with $\mu_{[k]}$.

Remark 5: The procedure (2.8) never requires more than $W_\lambda+1$ stages to terminate experimentation.

Remark 6: The procedure (2.8) permanently eliminates apparently non-contending populations; thus the number of observations taken at the m th stage of experimentation is less than or equal to the number of observations taken at the $(m-1)$ st stage of experimentation.

Remark 7: The cost of experimentation using procedures (2.7) and (2.8) can be measured in terms of expected number of stages to terminate experimentation and/or expected total number of observations to terminate experimentation. Which one is an appropriate measure will depend on the practical situation at hand.

Remark 8: Ramberg [1966] has demonstrated using Monte Carlo sampling methods that

$$\max_{\mu_1, \mu_2, \dots, \mu_k} E\{\text{Number of stages to terminate experimentation}\}$$

and

$$\max_{\mu_1, \mu_2, \dots, \mu_k} E\{\text{Total number of observations to terminate experimentation}\}$$

are less for (2.8) than for (2.7) when P^* is high (i.e., close to unity) but the inequality is reversed if P^* is sufficiently small; Perng [1969] has studied that question analytically. This result is of practical interest since it compares the performance of (2.7) and (2.8) when $\mu_{[1]} = \mu_{[k]}$, i.e., when, unknown to the experimenter, all of the population means are equal and thus the expected number of stages and the expected total number of observations are at their maxima.

Remark 9: Fabian [1974] pointed out the advantage of choosing $\lambda = \delta^*/2$, and recommended for that choice of λ that $1-P^*$ in a_λ be replaced by $2(1-P^*)$ yielding $a'_{\delta^*/2} = [2\sigma^2/\delta^*] \log[(k-1)/2(1-P^*)] = b$ (say); then b replaces a_λ and $\delta^*/2$ replaces λ in (2.8). This modified procedure still guarantees the probability requirement (2.2) when the population means are in the LF-configuration. It uniformly (in the μ_i) reduces the expected number of stages and expected total number of observations relative to the ones that would have been obtained with the unmodified procedure employing $\lambda = \delta^*/2$; in addition, in either the family of unmodified Paulson procedures or in the family of modified Paulson procedures the choice $\lambda = \delta^*/2$ has the property that

$$\max_{\mu_1, \mu_2, \dots, \mu_k} E\{\text{Total number of observations to terminate experimentation}\} \text{ is}$$

approximately minimized for P^* close to unity.

2.4.4 Two-stage procedure

The sequential procedures (2.7) and (2.8) have the drawbacks that they may not be appropriate for use in certain types of experimentation. For example, in agricultural experimentation where yields can be obtained only once per year (or per growing season), and thus only one vector of observations can be obtained per time period, multi-stage experimentation is impractical.

In such situations two-stage experimentation would appear to be appropriate. Alam [1970] and Tamhane [1975] have developed two-stage procedures which guarantee the probability requirement (2.2) when the population means are in the LF-configuration; their procedures screen out the apparently non-contending populations in the first stage, and concentrate sampling on the remaining populations in the second (terminal) stage. Tamhane's procedure has the added virtue of possessing a minimax property similar to that achieved by Fabian's modification of (2.8) when $\lambda = \delta^*/2$.

2.5 Procedures for use with the indifference-zone approach under the assumption of common unknown or completely unknown variances

As was mentioned in Section 2.3.2, if the experimenter wishes to guarantee (2.2) and adopts assumption (2.5c) or (2.5d) then he cannot use a single-stage procedure. In this section we shall consider two-stage procedures which accomplish these objectives.

2.5.1 Two-stage procedure for the common unknown variance case

The following two-stage procedure for the common unknown variance case was proposed by Bechhofer, Dunnett, and Sobel [1954]; constants $h_{k,P^*,n}$ (see c), below) necessary to implement this procedure for $P^* = 0.95$ are given in Table II.

- "a) In the first stage take an arbitrary common number $N_0 > 1$ of observations from each of the k populations.
- b) Calculate $S^2 = \sum_{i=1}^k \sum_{j=1}^{N_0} (X_{ij} - \sum_{j=1}^{N_0} X_{ij}/N_0)^2/n$ which is an unbiased estimate of σ^2 based on $n = k(N_0-1)$ degrees of freedom.
- c) Enter the appropriate table (e.g., Table II, below, for $P^* = 0.95$) with $n = k(N_0-1)$ and the specified P^* , and obtain a constant $h_{k,P^*,n} = h$ (say).
- d) In the second stage, take a common number $N-N_0$ of additional observations from each of the k populations where

$$N = N_0 \quad \text{if } 2(hS/\delta^*)^2 < N_0$$

$$N = [2(hS/\delta^*)^2] \quad \text{if } 2(hS/\delta^*)^2 > N_0.$$

(2.9)

and $[y]$ denotes the smallest integer equal to or greater than y .

e) Calculate the k over-all (first-stage plus second stage) sample

sums $\sum_{j=1}^N X_{ij} \quad (1 \leq i \leq k)$, and let $\sum_{j=1}^N X_{[1]j} < \sum_{j=1}^N X_{[2]j} < \dots <$

$\sum_{j=1}^N X_{[k]j}$ denote the ranked values of the $\sum_{j=1}^N X_{ij}$.

f) Select the population which yielded $\sum_{j=1}^N X_{[k]j}$ as the one associated with $\mu_{[k]}$.

Note: The constants $h_{k,p^*,n}$ are computed under the assumption that the $\mu_i \quad (1 \leq i \leq k)$ are in the LF-configuration.

Table II

Values of $h_{k,p^*,n}$ for $p^* = 0.95$

n	k					
	2	3	4	5	7	10
5	2.02	2.44	2.68	2.85	3.08	3.30
6	1.94	2.34	2.56	2.71	2.92	3.12
7	1.89	2.27	2.48	2.62	2.82	3.01
8	1.86	2.22	2.42	2.55	2.74	2.92
9	1.83	2.18	2.37	2.50	2.68	2.86
10	1.81	2.15	2.34	2.47	2.64	2.81
15	1.75	2.07	2.24	2.36	2.51	2.67
20	1.72	2.03	2.19	2.30	2.46	2.60
30	1.70	1.99	2.15	2.25	2.40	2.54
60	1.67	1.95	2.10	2.21	2.35	2.48
∞	1.64	1.92	2.06	2.16	2.29	2.42

The values in this table are abstracted from Table 1a of Dunnett [1955]; Table 1b of Dunnett [1955] gives corresponding values for $P^* = 0.99$; Dunnett's p equals our $k-1$.

Note: The value of $h_{k,P^*,n}$ given for the $n = \infty$ row of Dunnett [1955], Table 1a, is the same as the value given by Gupta [1963], Table I, for the same $k-1 = p = n$ and $P^* = 0.95 = 1-\alpha$.

Remark 10: The total number of observations N required by the two-stage procedure is a random variable since its value depends on the value of S^2 ; no additional observations are taken in the second stage if S^2 is sufficiently small.

Remark 11: Paulson [1964], Section 5, proposed an open-ended sequential procedure which permanently eliminates non-contending populations; his procedure is applicable in situations in which the common variance is unknown.

2.5.2 Two-stage procedures for the completely unknown variance case

Dudewicz and Dalal [1971], and also Rinott [1974], proposed two-stage procedures for the completely unknown variance case. Like (2.9), the common number of observations in the first stage for each of these procedures is arbitrary (>1), while the total number of observations per population is a random variable.

2.6 Procedure for use with the subset approach under the assumption of common (known or unknown) variance

As was mentioned in Section 2.3.2, if the experimenter wishes to guarantee (2.4) and adopts assumption 2.5a) or 2.5c), then he can use a single-stage procedure. The following single-stage procedure was proposed by Gupta [1956], [1965] for use under assumption 2.5c); constants $d_{k,P^*,n}$ (see c), below) necessary to implement this procedure are given in Table III. (Under assumption 2.5a, the random variable S in d) of (2.10) is replaced by σ , and the value of $d_{k,P^*,n}$ for $n = \infty$ is used.)

"a) Take a common arbitrary number $N > 1$ of observations from each of the k populations.

b) Calculate $\bar{X}_i = \sum_{j=1}^N X_{ij}/N$ ($1 \leq i \leq k$) and let $\bar{X}_{[1]} < \bar{X}_{[2]} < \dots < \bar{X}_{[k]}$ denote the ranked values of the \bar{X}_i ; also calculate

$S^2 = \sum_{i=1}^k \sum_{j=1}^N (X_{ij} - \sum_{j=1}^N X_{ij}/N)^2/n$ which is an unbiased estimate of σ^2 based on $n = k(N-1)$ degrees of freedom.

(2.10)

c) Enter the appropriate table (e.g., Table III, below, for $P^* = 0.95$) with $n = k(N-1)$ and the specified P^* , and obtain a constant $d_{k,P^*,n} = d$ (say).

d) Retain the population Π_i ($1 \leq i \leq k$) in the selected subset if and only if $\bar{X}_i \geq \bar{X}_{[k]} - dS/\sqrt{N}$.

Table III

Values of $d_{k,P^*,n}$ for $P^* = 0.95$

n	k		
	2	5	10
15	2.48	3.34	3.78
20	2.44	3.25	3.67
30	2.40	3.19	3.59
60	2.36	3.12	3.50

The values in this table are abstracted from Table I of Gupta and Sobel [1957] which gives many additional d-values for $P^* = 0.75, 0.90, 0.95, 0.975, 0.99$.

Note: $d_{k,P^*,n} = \sqrt{2} h_{k,P^*,n}$ where $h_{k,P^*,n}$ is given in Table II.

Remark 12: The width of the "yardstick" in d) of (2.10) is dS/\sqrt{N} which decreases with N ; thus the larger the value of N , the smaller the expected number of populations that will be included in the selected subset. Also, for fixed N

the more favorable the configuration of the population means (e.g., the larger the value of $(\mu_{[k]} - \mu_{[k-1]})/\sigma$), the smaller the expected number of populations that will be included in the selected subset. (This expected number always lies between unity and kP^* .)

Remark 13: In practice the subset approach is often used for screening purposes, since it tends to eliminate "non-contending" populations (i.e., those with small μ -values) from the selected subset. The populations retained in the subset can then be subjected to further study in an independent follow-up experiment in which the indifference-zone approach (say) is used.

2.7 Factorial experiments involving means

The statistical model given in Section 2.1 is appropriate for single-factor experiments. In a two-factor experiment we have rc normal populations Π_{ij} ($1 \leq i \leq r, 1 \leq j \leq c$) with population means μ_{ij} and population variances σ_{ij}^2 . It is sometimes appropriate to assume that $\mu_{ij} = \mu + \alpha_i + \beta_j$ ($\sum_{i=1}^r \alpha_i = \sum_{j=1}^c \beta_j = 0$), i.e., that there is no interaction between the factors, and that $\sigma_{ij}^2 = \sigma^2$ ($1 \leq i \leq r, 1 \leq j \leq c$). Here the α_i and the β_j are referred to as the "effects" of the first and second factor, respectively. It is assumed that μ , the α_i , the β_j , and σ^2 are unknown. Let $\alpha_{[1]} \leq \alpha_{[2]} \leq \dots \leq \alpha_{[r]}$ and $\beta_{[1]} \leq \beta_{[2]} \leq \dots \leq \beta_{[c]}$ denote the ranked values of the α_i and the β_j ; it is assumed that the pairing of the Π_{ij} with the $\alpha_{[i]}$ and $\beta_{[j]}$ ($1 \leq i \leq r, 1 \leq j \leq c$) is completely unknown.

In the above setup it is possible to consider goals such as

Goal: "To select the 'level' of the first factor associated with $\alpha_{[r]}$, and simultaneously to select the 'level' of the second factor associated with $\beta_{[c]}$," (2.11)

with associated probability requirements. Such problems are treated for the indifference-zone approach in Section 4 of Bechhofer [1954]. The virtue of conducting factorial experiments in this situation is discussed by Bawa [1972]. The indifference-zone selection procedures of Sections 2.4 and 2.5 can be used in multi-factor experiment; it is only necessary to make appropriate modifications in the procedures.

It is also possible to conduct single-factor or multi-factor ranking and selection experiments using the standard experimental designs such as randomized blocks and Latin squares, and these designs play the same type of role here as they do in classical hypothesis-testing situations.

2.8 Means vs. a fixed known standard

In Section 2.4.1-2.4.4 and 2.5.1-2.5.2 the selection procedures proposed were devised to select the category associated with the largest μ -value. However, in certain classes of experiments even the "best" one of the competing categories, i.e., the category with the largest μ -value, may not be good enough to warrant the experimenter's selecting it. For example, if the competing categories are drugs, the best one may not be worthy of consideration unless the expected period of immunity obtained with that drug is at least some specified period of time; or if the competing categories are types of heat treatment of steel, the best one may not be deemed satisfactory unless the expected tensile strength resulting from that type of treatment is at least some specified minimum value. Such types of problems involving comparisons of means with a fixed known standard are considered by Bechhofer and Turnbull [1974], [1975a]; in the first paper a single-stage procedure is proposed under assumption (2.5a), and in the second a two-stage procedure is proposed under assumption (2.5c). These procedures are generalizations of Bechhofer [1954] and Bechhofer, Dunnett, and Sobel [1954]. Gupta and Sobel [1958] proposed a single-stage procedure for this problem using the subset approach.

3. The normal variances problem

Section 2 dealt with the normal means problem. Corresponding procedures exist for the normal variances problem. Ranking and selection problems involving variances arise, for example, when the ordnance engineer is interested in selecting that type of projectile which yields the smallest dispersion of range, or when the laboratory technician is interested in selecting that measuring instrument which has the highest precision (e.g., that scale which has the greatest reproducibility). An analogue of the single-stage procedure given in Bechhofer [1954] for normal means is given in Bechhofer and Sobel [1954] for normal variances; factorial experiments involving variances are treated in Bechhofer [1968a] and [1968b] using a model proposed in Bechhofer [1960].

Bechhofer and Turnbull [1975b] is the counterpart for variances of Bechhofer and Turnbull [1974]. An analogue of the procedure given in Gupta [1956] for normal means is given in Gupta and Sobel [1962] for normal variances.

4. The Bernoulli p problem, and other problems

Ranking and selection problems involving Bernoulli p's (i.e., probabilities of "success" on a single trial) arise, for example, when a consumer is interested in selecting that producer whose product has the smallest fraction defective. An analogue of the procedure given in Bechhofer [1954] for normal means is given in Huyett and Sobel [1957] for Bernoulli p's, while the counterpart of the procedure given in Gupta [1956] for normal means is given in Gupta and Sobel [1960] for Bernoulli p's.

Sobel [1954] proposed a sequential procedure for selecting the exponential population with the largest mean; his results have applicability in reliability studies. Bechhofer, Kiefer, and Sobel [1968], p. 63 considered sequential procedures for ranking parameters of certain stochastic processes such as the Poisson process and the Wiener process. Various research workers have proposed procedures for many other ranking and selection problems involving parameters of distributions arising in practice.

5. Closing remarks

The ranking and selection formulation of statistical problems involving inferences concerning $k \geq 2$ categories has wide applicability in the solution of problems arising in experimentation. In this paper we have sketched only a small number of the relevant ideas and procedures. The interested reader is referred to Bechhofer, Kiefer, and Sobel [1968] for references up to that date, and to Gupta and Panchapakesan [1972] for references to the latter date concerning the subset approach. Additional and more recent references are given by Wetherill and Ofosu [1974]. The writer would appreciate learning of experimental situations in which some of the procedures described herein have proved helpful.

6. Acknowledgment

The writer would like to express his appreciation to the Army Research Office-Durham and the Office of Naval Research which have generously supported his research and that of his colleagues at Cornell who have been working with him on this program.

7. References

- Alam, K. (1970): "A two-sample procedure for selecting the population with the largest mean from k normal populations," Annals of the Institute of Statistical Mathematics (Tokyo), Vol. 22, pp. 127-136.
- Bawa, V. S. (1972): "Asymptotic efficiency of one R-factor experiment relative to R one-factor experiments for selecting the best normal population," Journal of the American Statistical Association, Vol. 67, No. 339, pp. 660-661.
- Bechhofer, R. E. (1954): "A single-sample multiple-decision procedure for ranking means of normal populations with known variances," Annals of Mathematical Statistics, Vol. 25, No. 1, pp. 16-39.
- Bechhofer, R. E. (1960): "A multiplicative model for analyzing variances which are affected by several factors," Journal of the American Statistical Association, Vol. 55, No. 290, pp. 245-264.
- Bechhofer, R. E. (1968a): "Single-stage procedures for ranking multiply-classified variances of normal populations," Technometrics, Vol. 10, No. 40, pp. 693-714.
- Bechhofer, R. E. (1968b): "Designing factorial experiments to rank variances," Transactions of the Twenty-Second Annual Conference of the American Society for Quality Control, pp. 69-73.
- Bechhofer, R. E., Dunnett, C. W. and Sobel, M. (1954): "A two-sample multiple decision procedure for ranking means of normal populations with a common unknown variance," Biometrika, Vol. 41, Parts 1 and 2, pp. 170-176.
- Bechhofer, R. E., Kiefer, J. and Sobel, M. (1965): Sequential Identification and Ranking Procedures (with special reference to Koopman-Darmois populations). University of Chicago Press.
- Bechhofer, R. E. and Sobel M. (1954): "A single-sample multiple-decision procedure for ranking variances of normal populations," Annals of Mathematical Statistics, Vol. 25, No. 2, pp. 273-289.
- Bechhofer, R. E. and Turnbull, B. W. (1974): "A $(k+1)$ -decision single-stage selection procedure for comparing k normal means with a fixed known standard: the case of common known variance," Technical Report No. 242, Department of Operations Research, Cornell University. Submitted for publication.
- Bechhofer, R. E. and Turnbull, B. W. (1975a): "A $(k+1)$ -decision two-stage selection procedure for comparing k normal means with a fixed known standard: the case of common unknown variance." In preparation.
- Bechhofer, R. E. and Turnbull, B. W. (1975b): "A $(k+1)$ -decision single-stage selection procedure for comparing k normal variances with a fixed known standard." In preparation.
- Deely, J. J. and Gupta (1968): "On the properties of subset selection procedures," Sankhyā, A, Vol. 30, Part 1, pp. 37-49.

- Dudewicz, E. J. (1971): "Nonexistence of a single-sample selection procedure whose $P(CS)$ is independent of the variances," South African Statistical Journal, Vol. 5, pp. 37-39.
- Dudewicz, E. J. and Dalal, S. R. (1971): "Allocation of observations in ranking and selection with unequal variances," a paper presented at the Symposium on Optimizing Methods in Statistics held at the Ohio State University, June 14-16, 1971. Accepted for publication in *Sankhyā*.
- Dunnett, C. W. (1955): "A multiple comparison procedure for comparing several treatments with a control," Journal of the American Statistical Association, Vol. 50, No. 272, pp. 1096-1121.
- Dunnett, C. W. (1960): "On selecting the largest of k normal population means," Journal of the Royal Statistical Society, B, Vol. 22, No. 1, pp. 1-40.
- Fabian, V. (1974): "Note on Anderson's sequential procedures with triangular boundary," Annals of Statistics, Vol. 2, No. 1, pp. 170-176.
- Fairweather, W. R. (1968): "Some extensions of Somerville's procedure for ranking means of normal populations," Biometrika, Vol. 55, No. 2, pp. 411-418.
- Gupta, S. S. (1956): "On a decision rule for a problem in ranking means," Institute of Statistics Mimeograph Series No. 150, University of North Carolina, Chapel Hill, N.C.
- Gupta, S. S. (1963): "Probability integrals of multivariate normal and multivariate t ," Annals of Mathematical Statistics, Vol. 34, No. 3, pp. 792-828.
- Gupta, S. S. (1965): "On some multiple decision (selection and ranking) rules," Technometrics, Vol. 7, No. 25, pp. 225-245.
- Gupta, S. S. and Panchapakesan, S. (1972): "On multiple decision procedures," Journal of Mathematical and Physical Sciences, Vol. VI, No. 1, pp. 1-72.
- Gupta, S. S. and Sobel, M. (1957): "On a statistic which arises in selection and ranking problems," Annals of Mathematical Statistics, Vol. 28, No. 4, pp. 957-967.
- Gupta, S. S. and Sobel, M. (1958): "On selecting a subset which contains all populations better than a standard," Annals of Mathematical Statistics, Vol. 29, No. 1, pp. 235-244.
- Gupta, S. S. and Sobel, M. (1960): "Selecting a subset containing the best of several binomial populations," Contributions to Probability and Statistics (ed. by Olkin et al), Stanford University Press, Stanford, California, pp. 1-25.
- Gupta, S. S. and Sobel, M. (1962): "On selecting a subset containing the population with the smallest variance," Biometrika, Vol. 49, Parts 3 and 4, pp. 495-507.

- Hall, W. J. (1959): "The most economical character of some Bechhofer and Sobel decision rules," Annals of Mathematical Statistics, Vol. 30, No. 4, pp. 964-969.
- Huyett, M. J. and Sobel, M. (1957): "Selecting the best one of several binomial populations," The Bell System Technical Journal, Vol. 36, pp. 537-576.
- Mahamunulu, D. M. (1967): "Some fixed-sample ranking and selection problems," Annals of Mathematical Statistics, Vol. 38, No. 4, pp. 1079-1091.
- Paulson, E. (1964): "A sequential procedure for selecting the population with the largest mean from k normal populations," Annals of Mathematical Statistics, Vol. 35, No. 1, pp. 174-180.
- Perng, S. K. (1969): "A comparison of the asymptotic expected sample sizes of two sequential procedures for ranking problem," Annals of Mathematical Statistics, Vol. 40, No. 6, pp. 619-632.
- Raiffa, H. and Schlaifer, R. (1961): "Selection of the best of several processes," Chapter 5B of Applied Statistical Decision Theory, Graduate School of Business Administration, Harvard University.
- Ramberg, J. S. (1966): "A comparison of the performance characteristics of two sequential procedures for ranking the means of normal populations," Technical Report No. 4, Department of Operations Research, Cornell University.
- Rinott, Y. (1974): "On two stage procedures for selecting the population with the largest mean from several normal populations with unknown variances," Report, Department of Mathematics, Cornell University. Submitted for publication.
- Santner, T. J. (1975): "A restricted subset selection approach to ranking and selection problems." Accepted for publication in the Annals of Statistics.
- Sobel, M. (1954-5): "Sequential procedures for selecting the best exponential population," Proceedings of the Third Berkeley Symposium on Mathematical Statistics and Probability, Vol. 5, pp. 99-110.
- Somerville, P. N. (1954): "Some problems of optimum sampling," Biometrika, Vol. 41, Parts 3 and 4, pp. 420-429.
- Tamhane, A. (1975): "A two-stage minimax procedure for selecting the population with the largest mean from k normal populations." Ph.D. dissertation (in preparation), Department of Operations Research, Cornell University.
- Wetherill, G.B. and Ofosu, J.B. (1974): "Selection of the best of k normal populations," Journal of the Royal Statistical Society, C, Vol. 23, No. 3, pp. 253-277.